مقدمه

- آرشیو اینترنت سازمانی غیر انتفاعی است که یک کتابخانه دیجیتال از سایت‌های اینترنت و دیگر مصنوعات فرهنگی به شکل دیجیتال را ساخته است.
- در ماه می سال ۱۹۹۶ تاسیس شده است.
- ماموریتش: فراهم کردن دسترسی جهانی به تمام دانش.
- با wayback machine می‌توان به اطلاعاتی که بیش از ۲۰ سال پیش منتشر می‌شده است، دسترسی پیدا کرد.
- از طریق برنامه Achive-It صفحات وب مهم، شناسایی می‌شوند.

امروزه آرشیو اینترنت شامل:

- ۲۷۹ میلیارد صفحات وب
- ۱۱ میلیون کتاب و متن
- ۴ میلیون صداهای ضبط شده (شامل ۱۶۰۰۰۰ کنسرت‌های زنده)
- ۳ میلیون ویدئو (شامل ۱ میلیون برنامه‌های خبری تلویزیون)
- ۱ میلیون عکس
- ۱۰۰۰۰۰ برنامه های نرم افزاری می باشد.

- هر کسی در Internet Archive می تواند رسانه ها را با free account بارگذاری کند.
- در سال ۲۰۰۵ دیجیتال‌سازی کتاب‌ها را شروع کرده است.
- امروزه در ۲۸ محل جهان ۱۰۰۰ کتاب در روز را اسکن می‌کنند
- کتاب‌هایی که قبل از ۱۹۲۳ منتشر شده‌اند نیز برای دانلود قابل دسترس هستند و صدها کتاب از هزاران کتاب مدرن از طریق سایت open library به امانت داده می‌شوند
- در اواخر سال ۲۰۰۰ شروع به بایگانی کردن برنامه‌های تلویزیونی کردند.
- اطلاعات پروژه‌های Internet Archive را می‌توان از طریق Blog (شامل آگهی‌های مهم)، contact us، خریدن مخزن و به دنبال آن Twitter و Facebook موسسه، به دست آورد.
- حمایت‌های مالی موسسه توسط افراد و سازمان‌های مختلفی تامین می‌شود.

چگونه می‌توانیم با آرشیو اینترنت ارتباط داشته باشیم:

خیابان ۳۰۰ فانستون، سان فرانسیسکو، CA 94118

شماره تلفن: ۴۱۵۵۶۱۶۷۶۷

فکس: ۴۱۵۸۴۰۰۳۹۱

صفحه وب: archive.org

پست الکترونیکی: info@archive.org

- توسعه سایر کتابخانه‌های اینترنتی از طریق بحث، گفت‌وگو و ابزارهای دیگر ترویج می‌یابد.
- می‌توان از طریق BitCoin کمک‌های خود را به سازمان اهدا کرد.
- از آن‌جایی که internet archive سازمانی غیر انتفاعی است کمک‌های مالی را از طریق منابع مختلف دریافت می‌کند.
- با خدمات مرجع کتابخانه‌های محلی و دانشگاهی می‌توان به تحقیقات کمک کرد.
- آمارهای کاربران در قسمت views سایت قرار دارد (https://archive.org/stats/)
- دلیل اهمیت مجموعه archive به خاطر حفظ و نگه‌داری فرهنگ و میراث جهانی است.

نکات جست‌وجو

- Advanced search سایت در قسمت پایین فیلد جست‌وجو قرار دارد.
- با مجوز creative commons نیز می‌توان جست‌وجو را انجام داد.
- نتایج جست‌وجو را می‌توان با گزینه‌های sort by مرتب کرد.
- اگر بخواهیم فقط در یک مجموعه جست‌وجو کنیم در قسمت راست صفحه از گزینه‌ی search this collection استفاده می‌کنیم.
- در قسمت sort by می‌توان از نمایش لیست به جای نمایش عنوان استفاده کرد.
- در موتور جست‌وجو فقط metadata درون صفحه نمایه می‌شود. پس موتور جست و جو متادیتای متن کتاب‌ها، فایل فردی را ندارد.
- فیلم‌های Prelinger:

- این نوع فیلم‌ها در قالب اصلی خود نگه‌داری می‌شوند، و نیاز به بازرسی فیلم، تعمیر هر گونه آسیب فیزیکی و نظارت توسط یک اپراتور ماهر است که رنگ، کنتراست، سرعت و تصویر را کنترل می-کند. برای دیجیتال‌سازی فیلم‌ها از scanstation, Flash KPEG, Virtual Dub استفاده می-شود.
- برای دوباره استفاده کردن از این فیلم‌ها بهتر است prelinger archive را مطلع کرد تا مسئولان بتوانند سایت را بهبود بخشند و یک کپی از تولید خود را به پست الکترونیکی سایت ارسال کنید.
- کاربران تجاری برای دسترسی به فیلم‌ها می توانند با نماینده‌ی prelinger archive (Getty Images) برای فروش فیلم‌ها مستقیما تماس بگیرند.
- با مجوز creative commons می توان به فیلم‌های prelinger دسترسی داشت و در تمام نقاط جهان می‌توان آن را به اشتراک گذاشت، مبادله، توزیع مجدد، انتقال و کپی کرد.

The Wayback Machine

- با wayback machine می‌توان به صفحات قدیمی دسترسی پیدا کرد و به آن‌ها لینک داد.
- ایده اصلی آن در سال ۱۹۹۶ شروع زمانی که internet archive، آرشیو وب را آغاز کرد.
- ویژگی منحصر به فرد wayback machine امکان مرور اسناد وب در چندین دوره زمانی است.
- نتایج جست‌وجو براساس تعداد تصاویر در wayback و تعداد لینک‌های مرتبط با صفحه اصلی سایت رتبه‌بندی می‌شوند.
- بایگانی صفحات وب وابسته به دسترس عموم بودن آن صفحه برمی‌گردد. صفحاتی که برای ورود به رمز عبور نیاز دارند بایگانی نمی‌کند.
- صفحاتی از wayback machine خارج می‌شوند که یک درخواست ایمیل به internet archive ارسال شده باشد.
- سیستم‌های چت یا پیام-های شخصی را که به بولتن‌های usenet یا تابلوهای تبلیغاتی آنلاین دسترسی ندارند جمع آوری و بایگانی نمی‌کند.
- از save page now برای ذخیره صفحه خاص استفاده می‌شود.
- زمانی که تصاویر در سرورهای سایت، آرشیو نشود تصاویر ناقص یا خاکستری می‌شوند. بهترین روش برای دیدن همه‌ی فایل‌های نگه‌داری شده در سایت: http://web.archive.org/*/www.yoursite.com/*

- برای پیدا کردن سایت‌هایی که فقط URL آن را دارید می‌توان از Wayback Machine استفاده کرد با زدن دکمه ی browse history
- سایت خود را می‌توان در wayback machine وارد کرد و برای اطمینان از آن باید سایت شما در دایرکتوری‌های آنلاین گنجانده شود و سایت‌های مشابه/ مرتبط با آن به شما لینک دهند.
- برای آپلود کردن هر فایلی در internet archive باید در قالب‌های استاندارد سایت و در فرمت فشرده باشد تا برای عموم قابل دسترس باشد.
- از طریق پروتکل BitTorrrent بیش از۱,۴ میلیون نسخه بایگانی در دسترس است که شامل تقریبا یک پتابایت از مواد دامنه عمومی است.
- BitTorrent یک پروتکل به اشتراک‌گذاری فایل است.
- Torrent راهی مناسب است تا بسیاری از فایل‌ها و یا محتویات بزرگ را آپلود کرد.
- Internet archive مکان خوبی برای ارسال موسیقی زنده و آپلود آن نیست.
- بسیاری از فیلم‌ها و مجموعه‌ها دارای مجوز Creative Commons هستند. آپلود کنندگان ممکن است مشخص کنند که آیا آیتم دارای مجوز CC است یا خیر. اگر این کار را بکنند، لوگوی Creative Commons در سمت چپ صفحه جزئیات فیلم ظاهر خواهد شد.
- شما می‌توانید فیلم‌هایی را که متعلق به کپی رایت خود هستید و یا در دامنه عمومی قرار دارد را آپلود کنید.
- برای آپلود باید حق نسخه‌برداری فیلم را داشته باشید.
- بنابراین شما باید فرمت‌های فایل که به راحتی قابل دانلود و/ یا قابل پخش برای سایر کاربران سایت ارسال کنید. فایل‌های MPEG2 ساده ترین نوع فایل برای ما است.
- در فیلم شما می‌توانید زیرنویس‌ها را نیز اضافه کنید. با نام IDENTIFIER.srt و اگر زیرنویس بیش از یک زبان داشته باشد زبان نیز به نام فایل افزوده می‌شود مثلا IDENTIFIER_english.srt
- فایل‌های stream only فقط به صورت آنلاین استفاده می-شوند و قابل دانلود نیستند.
- برای دانلود فایل از کلیک کپی استفاده کنید یا راست کلیک کنید و گزینه ذخیره را بزنید.
- اگر می‌خواهید تمامی فایل‌ها را در یک فرمت دانلود کنید در گزینه‌های دانلود روی فرمتی که می-خواهید دانلود کنید کلیک کنید و اگر چندین فایل شما در یک فرمت باشد آن را به صورت zip دانلود کنید.

- برای دانلود تمام فایل‌ها و فراداده‌های منبع اصلی در گزینه‌های دانلود روی number originals کلیک کنید.
- با توجه به مسایل حقوقی، برخی از کتاب‌ها فقط در DAISY (سیستم اطلاعاتی قابل دسترسی دیجیتال) در دسترس هستند.
- بازی Arcade اینترنت مجموعه‌ای از بازی‌های شبیه‌سازی شده دهه ۱۹۷۰ است که می‌توان در مرورگر بازی کرد. هیچ پلاگینی برای اجرای بازی اینترنت وجود ندارد.

Open library

- یک پروژه‌ی باز با ماموریت داشتن یک وب برای هر کتاب چاپ شده می‌باشد. نرم افزار، داده ها و استنادات باز هستند.
- کتاب‌ها در open library وجود ندارند، ممکن است لینک کتاب‌هایی باشد که در archive.org وجود دارد یا لینک‌هایی که شما فکر می‌کنید می‌توان کتاب را پیدا کرد.

Archive-It

- یک سرویس کمکی است که اجازه می‌دهد موسسات، مجموعه‌های تولید شده با محتوای دیجیتالی را بسازند و نگه‌داری کنند.
- از طریق کاربرد کاربرپسندی وب، شریکان archive-It می‌توانند مجموعه-های بایگانی شده‌شان را جمع‌آوری کنند، فهرست‌بندی کنند، مدیریت کنند و قرض دهند.
- مجموعه‌های در مرکز داده‌های internet archive میزبانی می‌شوند و با جست‌وجو تمام متن برای عموم قابل دسترس‌اند.

مقوله‌های قابل جست‌وجو

صفحات وب

متون کتاب‌ها

منابع ویدئویی

منابع صوتی

برنامه‌های تلویزیونی

نرم افزارها

عکس‌ها

کنسرت‌ها

مجموعه‌ها